

METHOD

Open Access



# Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data

Francesca Finotello<sup>1</sup>, Clemens Mayer<sup>1</sup>, Christina Plattner<sup>1</sup>, Gerhard Laschober<sup>1</sup>, Dietmar Rieder<sup>1</sup>, Hubert Hackl<sup>1</sup>, Anne Krogsdam<sup>1</sup>, Zuzana Loncova<sup>1</sup>, Wilfried Posch<sup>2</sup>, Doris Wilflingseder<sup>2</sup>, Sieghart Sopper<sup>3</sup>, Marieke Ijsselsteijn<sup>4</sup>, Thomas P. Brouwer<sup>4</sup>, Douglas Johnson<sup>5,6</sup>, Yaomin Xu<sup>7</sup>, Yu Wang<sup>7</sup>, Melinda E. Sanders<sup>8</sup>, Monica V. Estrada<sup>8</sup>, Paula Ericsson-Gonzalez<sup>8</sup>, Pornpimol Charoentong<sup>9,10</sup>, Justin Balko<sup>5,6</sup>, Noel Filipe da Cunha Carvalho de Miranda<sup>4</sup> and Zlatko Trajanoski<sup>1,11\*</sup> 

## Abstract

We introduce quanTlseq, a method to quantify the fractions of ten immune cell types from bulk RNA-sequencing data. quanTlseq was extensively validated in blood and tumor samples using simulated, flow cytometry, and immunohistochemistry data.

quanTlseq analysis of 8000 tumor samples revealed that cytotoxic T cell infiltration is more strongly associated with the activation of the CXCR3/CXCL9 axis than with mutational load and that deconvolution-based cell scores have prognostic value in several solid cancers. Finally, we used quanTlseq to show how kinase inhibitors modulate the immune contexture and to reveal immune-cell types that underlie differential patients' responses to checkpoint blockers.

**Availability:** quanTlseq is available at <http://icbi.at/quantiseq>.

**Keywords:** Cancer immunology, Immunotherapy, Deconvolution, RNA-seq, Immune contexture

## Background

Cancer immunotherapy with antibodies targeting immune checkpoints has shown durable benefit and even curative potential in various cancers [1, 2]. As only a fraction of patients respond to immune checkpoint blockers, efforts are underway to identify predictive markers for cancer immunotherapy and mechanistic rationale for combination therapies. We have previously shown that the immune contexture—the type and density of tumor-infiltrating immune cells—has a prognostic value in colorectal cancer (CRC) [3]. Later, the association between the densities of tumor-infiltrating immune cells and patient overall survival was confirmed in different primary and metastatic cancers [4]. In particular, cytotoxic CD8<sup>+</sup> T cells, which can

specifically recognize and kill tumor cells, are associated with a good clinical outcome in different cancer types [5] and have a pivotal role in anti-PD1 immunotherapy [1]. Therefore, the quantification of the immune contexture of human tumors can not only unveil prognostic markers, but also provide relevant information for the prediction of response to checkpoint blockade.

Moreover, the quantification of the immune contexture of archived tumor samples holds the promise to identify drugs having additive or synergistic potential with immune checkpoint blockers. For example, since certain chemotherapeutic drugs induce immunogenic cell death [6], the analysis of a large number of samples could pinpoint patient subgroups that would benefit from the combination with immune checkpoint blockers. Similarly, as a number of targeted anticancer agents exhibit immunostimulatory activity [6], the quantification of the immune contexture could provide mechanistic rationale for the design of combination

\* Correspondence: [zlatko.trajanoski@i-med.ac.at](mailto:zlatko.trajanoski@i-med.ac.at)

<sup>1</sup>Biocenter, Division of Bioinformatics, Medical University of Innsbruck, Innrain 80, Innsbruck, Austria

<sup>11</sup>Austrian Drug Screening Institute, Innrain 66A, Innsbruck, Austria

Full list of author information is available at the end of the article



therapies. However, comprehensive and quantitative immunological characterization of tumors in a large number of clinical samples is currently hampered by the lack of simple and efficient methods. Cutting-edge technologies like single-cell RNA sequencing and multi-parametric flow or mass cytometry are technically and logistically challenging and cannot be applied to archived samples. Multiplexed immunohistochemistry (IHC) [7] or immunofluorescence (IF) assays can be performed only in specialized labs and require sophisticated equipment and extensive optimization of protocols for specific cancer entities. Moreover, manual and semi-automatic image analysis is required, which is highly time consuming and laborious. For an overview of imaging techniques for quantitative analysis of the tumor micro-environment, we refer to two recent reviews [8, 9].

Computational methods for quantitative immunophenotyping of tumors from bulk RNA sequencing (RNA-seq) data hold potential for efficient and low-cost profiling of a large number of samples, but currently suffer from several limitations. Bioinformatics methods based on immune-cell-specific markers like MCPcounter [10], xCell [11], or other approaches based on gene set enrichment analysis (GSEA) [12–14] compute only semi-quantitative scores that predict the enrichment of specific immune cell types in a sample, but that cannot be neither interpreted as cell fractions nor compared between cell types [15]. Deconvolution algorithms (reviewed in [16]) enable to quantitatively estimate the proportions of the cell types of interest. However, currently available deconvolution algorithms for immune cell quantification have several drawbacks [16]. For instance, CIBERSORT, a popular method based on support-vector regression for the deconvolution of 22 immune cell phenotypes, can only infer cell fractions relative to the total immune cell population and has been developed and validated using microarray data [17]. TIMER performs deconvolution of six immune cell types, but the results cannot be interpreted directly as cell fractions, nor compared across different immune cell types and data sets [18]. EPIC, a deconvolution method recently developed using RNA-seq data, estimates absolute fractions referred to the whole cell mixture, but does not consider immune cells relevant for cancer immunology like regulatory T cells ( $T_{reg}$ ) cells, dendritic cells, and classically (M1) and alternatively (M2) activated macrophages [19]. Hence, there is a need for a validated deconvolution-based algorithm that estimates absolute proportions of relevant immune cell types from RNA-seq data, thereby enabling inter-sample as well as intra-sample comparisons.

We therefore developed quanTIseq, a computational pipeline for the characterization of the tumor immune contexture using bulk RNA-seq data and imaging data from whole tissue slides. quanTIseq can quantify the absolute fractions of immune cells using a novel deconvolution approach and performs *in silico* multiplexed

immunodetection of the same cell types by integrating the deconvolution results with total cell densities extracted from images of IF, IHC, or hematoxylin and eosin (H&E)-stained tissue slides. We performed extensive validation using simulated data, published data sets, and *de novo* generated flow cytometry data. In addition, we validated quanTIseq using RNA-seq data and histological images from IHC/IF-stained slides from three independent cancer data sets. We then applied quanTIseq to analyze over 8000 solid tumors of The Cancer Genome Atlas (TCGA) [20] and show that the activation of the CXCR3/CXCL9 axis, rather than the mutational load, is associated with the infiltration of intratumoral cytotoxic T cells. Moreover, we observe highly heterogeneous immune contextures across and within tumors and show that the immunoscore and a T cell/B cell score computed from quanTIseq deconvolution results have prognostic values in several solid cancers. Finally, we demonstrate that the immune contexture of human tumors is pharmacologically modulated by kinase inhibitors and show that quanTIseq can be used to shed light on the features of the tumor immune contexture that underlie differential patients' responses to checkpoint blockade.

## Methods

### Collection of RNA-seq data from immune cell types and tumor cell lines

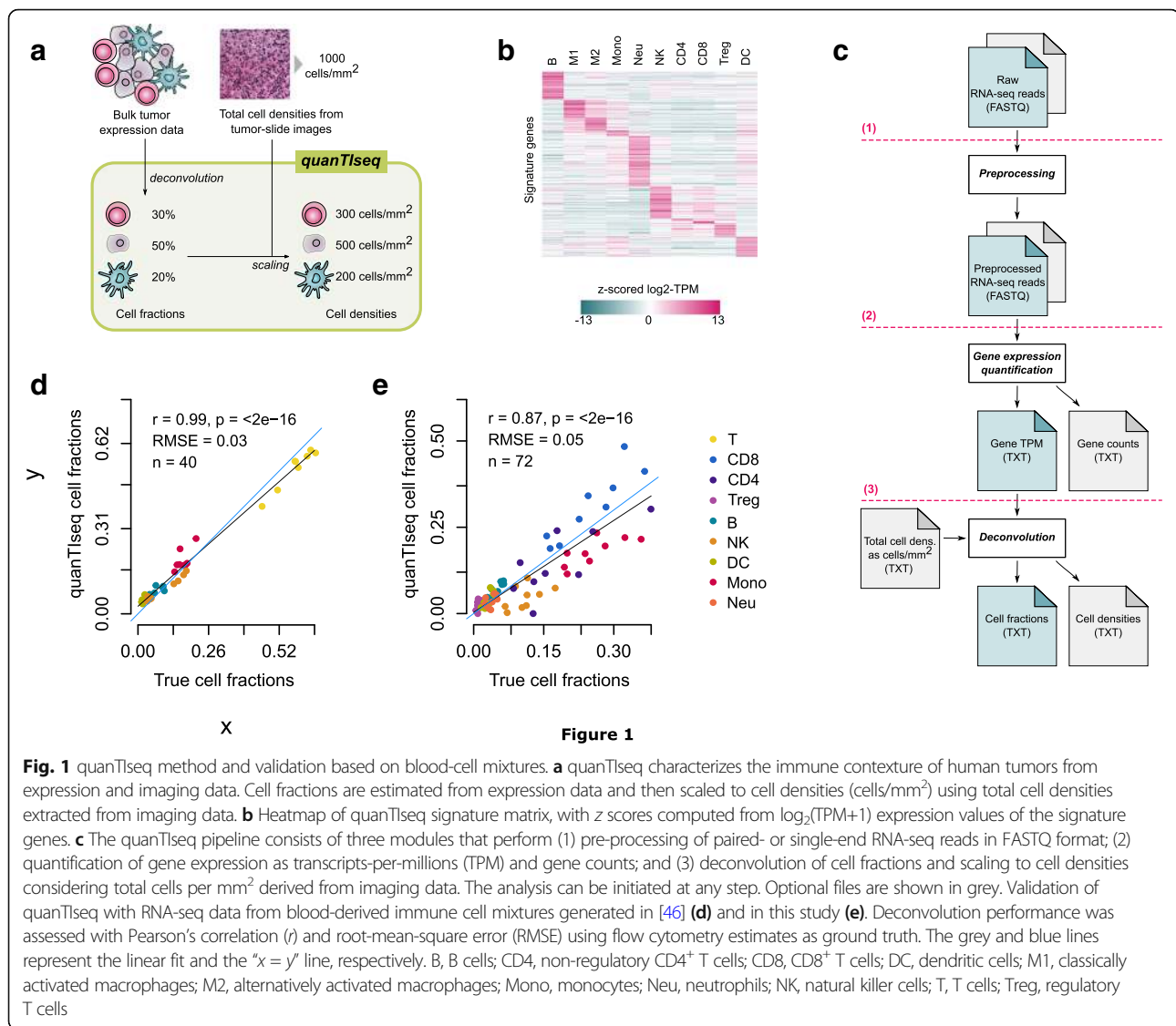
To build the signature matrix, we collected 51 data sets generated from paired-end Illumina RNA-seq of blood-derived immune cells (Additional file 1). In addition, we downloaded from the Cancer Genomics Hub (CGHub, accessed on February 2016) RNA-seq data from a breast (G41726.MCF7.5) and a colorectal (G27202.SW480.1) cancer cell line. BAM files of mapped reads gathered from the CGHub were converted to FASTQ with samtools [21], whereas SRA files downloaded from the Sequence Read Archive (SRA, <https://www.ncbi.nlm.nih.gov/sra/>) were converted to FASTQ with the “fastq-dump” function of the SRA Toolkit.

### RNA-seq data pre-processing

FASTQ files of RNA-seq reads were pre-processed with Trimmomatic [22] to remove adapter sequences and read ends with Phred quality scores lower than 20, to discard reads shorter than 36 bp, and to trim long reads to a maximum length of 50 bp. This analysis is implemented in the “Preprocessing” module of quanTIseq (step 1 in Fig. 1c), which also allows selecting different parameters for data preprocessing.

### Quantification of gene expression and normalization

The pre-processed RNA-seq reads were analyzed with Kallisto [23] to generate gene counts and transcripts per millions (TPM) using the “hg19\_M\_rCRS” human reference.



For single-end data, the following Kallisto options were used: “--single -l 50 -s 20”. After gene expression quantification, gene names were re-annotated to updated gene symbols defined by the HUGO Gene Nomenclature Committee (<http://www.genenames.org>, annotations downloaded on April 2017). In case of duplicates, the median expression per gene symbol was considered. The final expression value  $x_{gl}$  for each gene  $g$  in library  $l$  was computed from TPM with the following formula:

$$x_{gl} = \frac{TPM_{gl} \cdot 10^6}{\sum_i TPM_{il}} \quad (1)$$

For microarray data, before the normalization of Eq. 1, expression data were transformed from logarithmic to natural scale (when needed) and quantile-normalized. TPM can be computed from RNA-seq reads with the

“Gene Expression Quantification” module of quanTlseq (step 2 in Fig. 1c). Gene re-annotation and expression normalization are performed by the quanTlseq “Deconvolution” module before deconvolution (step 3 in Fig. 1c), and quantile normalization is performed if the “--arrays” option is set to “TRUE”.

#### Generation of the simulated data sets

We simulated RNA-seq data from breast tumors with different purity values and immune infiltrates by mixing pre-processed reads from immune cell types and from a tumor cell line (G41726.MCF7.5) of the RNA-seq compendium. We simulated 100 different immune cell mixtures by sampling the cell fractions from a uniform distribution in the [0–1] interval. The cell fractions were combined with 11 different tumor purity scenarios: 0:10:100% tumor purity, defined as the fraction of read

pairs from the tumor cell line over total read pairs. Each simulated data set consisted of one million paired-end reads. In addition, for the data set with 60% purity (which is the minimum value considered by the TCGA consortium for tumor specimen inclusion [24]), we simulated different sequencing depths, namely, 1, 2, 5, 10, 20, 50, and 100 million read pairs. In total, we generated 1700 simulated RNA-seq data sets.

### Generation of the TIL10 signature matrix

An expression matrix was generated from the compendium of RNA-seq data as described in “RNA-seq data pre-processing” and “Quantification of gene expression and normalization” and consisted in 19,423 genes and 53 immune and tumor cell libraries. From this matrix, we filtered out the genes that were not detected in at least two immune libraries and selected the genes specific for each cell type considering the criteria described in the following. Gene expression is here considered in terms of normalized values  $x_{gl}$  (Eq. 1) on a natural scale, if not differently stated.

### Cell-specific expression

We quantized the expression of each gene into three bins representing low, medium, and high expression, computed as in [25]. For each immune cell type, we selected the genes having (i) high quantized expression in all libraries belonging to the considered immune cell type and (ii) low or medium quantized expression in all other libraries.

### Expression in tumors

We filtered the signature genes that were highly expressed also in tumor cells by discarding the genes having a median  $\log_2$  expression larger than 7 in all non-hematopoietic cancer cell lines assayed in the Cancer Cell Line Encyclopedia (CCLE) [26], as done in [17]. Moreover, RNA-seq data from 8243 TCGA solid tumors were used to remove genes that provide little support for bulk-tissue deconvolution because their expression in tumor samples is generally low or null. More precisely, we discarded the genes having an average expression across all TCGA samples lower than 1 TPM.

### Specificity of marker genes

Since signature genes specific for a certain cell type should not be associated to another cell type, we considered a compendium of 489 gene sets specific for 64 cell types recently proposed in [11] and removed the signature genes that were listed in a gene set specific for another cell type. CD4<sup>+</sup> T cell gene sets were not used to filter T<sub>reg</sub> cell signature genes, as the CD4<sup>+</sup> T cell population may contain bona fide T<sub>reg</sub> cell expression markers such like the forkhead box P3 (FOXP3).

### Range of expression

As genes with high expression can bias deconvolution results, we excluded the genes whose expression exceeded 700 TPM.

### Correlation with true cell fractions

The 1700 simulated RNA-seq data sets (see the “[Generation of the simulated data sets](#)” section) were then used to identify the signature genes that provide valuable information over cell fractions and are more robust to the sequencing depth and unknown tumor content. For each cell type, we selected the genes whose expression levels had a correlation with the true cell fractions equal or greater than 0.6.

### Restricted expression

We considered four external expression data sets from enriched/purified immune cells: two microarray data sets (GEO accession: GSE28490 and GSE2849) [27], an RNA-seq data set [28], and a microarray compendium that was used to build the CIBERSORT LM22 signature matrix [17]. All data sets were preprocessed and normalized as explained in the previous paragraphs. For each gene  $g$  specific for a cell type  $c$  in the signature matrix, we computed the ratio  $R_{gd}$  between the median expression across all libraries in data set  $d$  belonging to the cell type  $c$  and the median expression across all libraries in data set  $d$  not belonging to the cell type  $c$ . For each cell type, the top 30 ranked signature genes (or less, when not available) with  $\text{median}_d(R_{gd}) \geq 2$  were selected for the final signature matrix. When processing the T<sub>reg</sub> signature genes, the data sets belonging to CD4<sup>+</sup> T cells were not considered. T<sub>reg</sub> signature genes were further filtered with a similar approach, but considering the RNA-seq data of circulating CD4<sup>+</sup> T and T<sub>reg</sub> cells from and selecting only the genes with  $\text{median}_d(R_{gd}) \geq 1$ .

The final signature matrix TIL10 (Additional file 1) was built considering the 170 genes satisfying all the criteria reported above. The expression profile of each cell type  $c$  was computed as the median of the expression values  $x_{gl}$  over all libraries belonging to that cell type:

$$x_{gc} = \text{median}_{l \in c}(x_{gl})$$

For the analysis of RNA-seq data, quanTIseq further reduces this signature matrix by removing a manually curated list of genes that showed a variable expression in the considered data sets: *CD36*, *CSTA*, *NRGN*, *C5AR2*, *CEP19*, *CYP4F3*, *DOCK5*, *HAL*, *LRRK2*, *LY96*, *NINJ2*, *PPP1R3B*, *TECPR2*, *TLR1*, *TLR4*, *TMEM154*, and *CD248*. This default signature considered by quanTIseq for the analysis of RNA-seq data consists of 153 genes and has a lower condition number than the full TIL10 signature (6.73 compared to 7.45), confirming its higher cell specificity. We advise using the full TIL10 matrix (--rmgenes="none") for the analysis of microarray data, as they often lack some

signature genes, and the reduced matrix (`--rmgenes= "default"`) for RNA-seq data. Alternatively, the `"rmgenes"` option allows specifying a custom list of signature genes to be disregarded (see `quanTIseq` manual).

### Deconvolution

The `quanTIseq` deconvolution module takes as input:

- A mixture matrix  $M_{gj}$  of expression values over  $g = 1, \dots, I$  genes and  $j = 1, \dots, J$  samples
- A signature matrix  $S_{gc}$  of expression values over  $g = 1, \dots, G$  signature genes and  $c = 1, \dots, C$  cell types

After re-annotation of gene symbols and normalization of the mixture matrix (see the ["Quantification of gene expression and normalization"](#) section), `quanTIseq` performs deconvolution of the unknown cell fractions  $F_{cj}$  over  $C$  immune cell types and  $J$  samples. For each sample  $j$ , the following system of equations is solved to estimate the cell fractions  $F_c$  (the subscript  $j$  is omitted):

$$M_{g|g \in G^*} = S_{g|g \in G^*} \times F_c$$

where  $G^*$  is the set of signature genes that are present in the mixture matrix. `quanTIseq` solves this inverse problem using constrained least squares regression, i.e., by minimizing the formula  $\|S \times F - M\|^2$ , imposing the constraints:

$$F_c \geq 0 \text{ for } c = 1, \dots, C$$

$$\sum_{c=1}^C F_c \leq 1$$

To account for the differences in the average mRNA content per cell type, which might otherwise bias deconvolution results [19, 29–31], the estimated cell fractions are normalized by a cell-type-specific scaling factor  $n_c$ :

$$F'_c = \frac{F_c}{n_c}$$

Then, the cell fractions are scaled so to sum up to the original percentage of total cells, as:

$$F''_c = \frac{F'_c \cdot f}{f'}$$

where

$$f = \sum_{c=1}^C F_c$$

$$f' = \sum_{c=1}^C F'_c$$

Finally, the proportion of "other" (uncharacterized) cells is estimated as:

$$F_{\text{other}} = 1 - \sum_{c=1}^C F''_c$$

As the population of other cells might include different types of malignant and normal cells with various mRNA contents [32] depending on the sample under investigation, `quanTIseq` does not scale these estimates. The scaling factors  $n_c$  were computed as the median expression of the Proteasome Subunit Beta 2 (PSMB2) housekeeping gene [33] across the immune cell types of the RNA-seq compendium and were highly correlated with experimentally-derived scaling factors used in the EPIC approach [19] (Pearson's correlation  $r = 0.86$  considering the immune cells in common). In the analysis of the simulated RNA-seq data, where the true fractions represented mRNA fractions and not cell fractions, deconvolution was performed without mRNA-content normalization (Additional file 2: Table S3).

The deconvolution of  $T_{\text{reg}}$  cells and  $CD4^+$  T cells is inherently hampered by the high correlation of their expression signatures (namely, *multi-collinearity* [17]) and can result in the underestimation of  $T_{\text{reg}}$  cells present in low fractions. Thus, we adopted a heuristic strategy to specifically address the issue of  $T_{\text{reg}}$  cell underestimation. First, `quanTIseq` estimates the  $T_{\text{reg}}$  cell fractions  $F_{\text{reg}}^1$  considering all cell types together. Then, for the samples with  $F_{\text{reg}}^1 < 0.02$ , `quanTIseq` re-estimates the  $T_{\text{reg}}$  cell fractions  $F_{\text{reg}}^2$  removing from the signature matrix the expression profiles of the  $CD4^+$  T cells. The final  $T_{\text{reg}}$  cell fractions are then estimated by averaging the results:

$$F_{\text{reg}} = \text{mean}(F_{\text{reg}}^1, F_{\text{reg}}^2)$$

whereas  $CD4^+$  T cell fractions are scaled to:

$$F_{\text{CD4}} = \max(F_{\text{CD4}}^1 - F_{\text{reg}}, 0)$$

Finally, all cell fractions are normalized to sum up to 1.

The analysis described in this section is implemented in the "Deconvolution" module of `quanTIseq` (step 3 in Fig. 1c).

The full `quanTIseq` pipeline can be applied to single or multiple samples and can be initiated at any step. For instance, pre-computed expression matrices can be analyzed directly with the deconvolution module (step 3 in Fig. 1c), although particular care must be taken when



performing data pre-processing and annotation of signature genes.

#### Deconvolution of bulk tumor expression data

Aberrant de-methylation and sequence duplication can lead to over-expression of immune signature genes. Tumor RNA-seq data can be analyzed with *quanTIseq* setting the “--tumor” option to “TRUE”. This setting discards the signature genes whose  $\log_2(x_{gl} + 1)$  expression in the TCGA RNA-seq data exceeds 11 TPM, which are *NUPR1*, *CD36*, *CSTA*, *HPGD*, *CFB*, *ECM1*, *FCGBP*, *PLTP*, *FXRD6*, *HOPX*, *SERPING1*, *ENPP2*, *GATM*, *PDPN*, *ADAM6*, *FCRLA*, and *SLC1A3*. All tumor data sets presented in this work have been analyzed with this parameter setting (Additional file 2: Table S3).

#### Publicly available validation data sets

To benchmark *quanTIseq*, we considered the expression data sets listed in Additional file 2: Table S1, using the options reported in Additional file 2: Table S3. Normalized microarray data were downloaded from the Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo>) with the GEOquery R package [34]. Probes were mapped to gene symbols with the biomaRt R package [35]. In case of multiple probes mapping to the same gene symbol, the probe with the highest average expression across all samples was selected. Immune cell fractions estimated with flow cytometry, Coulter Counter, or from images of stained tissue slides were used as ground truth to validate *quanTIseq*. Where necessary, different functional states of an immune cell type were aggregated by summing up the corresponding cell fractions (e.g., for the Newman’s data set [17], B cells were quantified summing up the fractions of naïve and memory B cells).

#### Generation of flow cytometry and RNA-seq data from blood-derived immune cell mixtures

Blood samples from healthy human donors were obtained from the Blood Bank Innsbruck under approval of the local ethics committee. Peripheral blood mononuclear cells (PBMC) were isolated from human whole blood by density centrifugation using Lymphocyte Separation Medium (Capricorn, Ebsdorfergrund, Germany). The PBMC fraction was collected and washed three times with Dulbecco’s phosphate buffered saline. To isolate polymorphonuclear (PMN) cells, the cells on top of the erythrocytes were collected and contaminating red blood cells were removed by two rounds of lysis with 0.2% NaCl solution at 4°C. PMN were added to the PBMC fractions in low abundance (3–6% of total cells), and aliquots were taken for RNA extraction and flow cytometry analysis. Total RNA was extracted with the Qia-gen RNeasy mini kit (Qiagen GmbH, Hilden, Austria), including on-column DNase I treatment. INVIEW

polyA RNA library preparation, and Illumina 50 bp SR sequencing at >60 Million reads per library, was obtained from an external provider (GATC Biotech, Konstanz, Germany).

The fractions of the following cell types in the immune cell mixtures were determined by flow cytometry using specific marker combinations: CD4<sup>+</sup> T cells (CD3<sup>+</sup>CD4<sup>+</sup>), CD8<sup>+</sup> T cells (CD3<sup>+</sup>CD8<sup>+</sup>), T<sub>reg</sub> cells (CD3<sup>+</sup>CD4<sup>+</sup>CD25<sup>+</sup>CD127<sup>-</sup>), B cells (CD19<sup>+</sup>), NK cells (CD3<sup>-</sup>CD16<sup>+</sup>CD56<sup>+</sup>), myeloid dendritic cells (Lin<sup>-</sup>HLA-DR<sup>+</sup>CD11c<sup>+</sup>), monocytes (CD14<sup>+</sup>), and neutrophils (CD15<sup>+</sup>CD16<sup>+</sup>). Labeled antibodies specific for the following antigens were purchased from BD Biosciences (San Jose, CA, USA) and Biolegend (San Diego, CA, USA): CD3 (UCHT1), CD4 (RPA-T4), CD8 (HIT8a), CD11c (3.9), CD14 (M5E2), CD15 (W6D3), CD16 (3G8), CD19 (HIB19), CD20 (2H7), CD25 (BC96), CD56 (B159), CD127 (A019D5), HLA-DR (L243), Lin: CD3, CD14, CD19, CD20, CD56. The measurements were performed on a BD LSRFortessa flow cytometer, and the data were evaluated with FlowLogic 7.1 software (Inivai Technologies, Melbourne, Australia).

#### Leiden validation data set

Fresh frozen and formalin-fixed material was available from 19 colorectal cancer patients (Additional file 3). Their usage was approved by the local ethics committee (P15.282). All the specimens were anonymized and handled according to the ethical guidelines described in the Code for Proper Secondary Use of Human Tissue in the Netherlands of the Dutch Federation of Medical Scientific Societies. RNA was isolated with the NucleoSpin RNA kit (Macherey-Nagel, Düren, Germany) including on-column DNase I treatment. Library preparation was preceded by rRNA depletion with the NEBNext rRNA depletion kit (New England Biolabs, MA, USA). PE 150 bp sequencing was performed at GenomeScan (Leiden, The Netherlands) on a HiSeq 4000 (Illumina, San Diego, CA, USA).

Four-micrometer sections of formalin-fixed paraffin-embedded tissues were deparaffinized and underwent heat-mediated antigen retrieval in 10 mmol/L citrate buffer solution (pH 6). Unspecific antibody binding was prevented with the SuperBlock PBS buffer (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer’s instructions. Immunofluorescence detection was performed using two panels. Firstly, the T cell panel contains the following antibodies: pan-cytokeratin (AE1/AE3, ThermoFisher scientific and C11, Cell Signalling Technology), anti-CD3 (D7A6E), and anti-CD8 (4B11, DAKO). Secondly, the myeloid panel contains the following antibodies: pan-cytokeratin (AE1/AE3, Novusbio and C11, Biolegend), anti-HLA-DR (TAL1B5, Thermo Fisher Scientific), anti-CD68 (D4B9C, Cell Signalling Technology), and anti-CD163 (10D6, Thermo Fisher Scientific). Immunofluorescent detection was performed directly and indirectly

with Alexa488, Alexa594, Alexa647, Alexa680, CF555, and CF633 using an in-house methodology [36].

For immunohistochemical detection, 4- $\mu$ m sections were deparaffinized after which endogenous peroxidase was blocked with a 0.3% hydrogen peroxide/methanol solution. Following heat-mediated antigen retrieval in 10 mmol/L citrate buffer solution (pH 6), overnight labeling was performed with anti-CD4 (EPR68551, Abcam), anti-FOXP3 (236A/E7), and CD20 (L26, Dako) respectively. After washing in PBS, Tissue sections were incubated for 1 h with Poly-horseradish peroxidase solution (Immunologic Duiven, The Netherlands) at room temperature. The slides were developed with the DAB+ chromogen (DAKO, Agilent Technologies, Santa Clara, CA, USA) solution and counterstained with hematoxylin (Thermo Fisher Scientific).

Image analysis for both immunofluorescence and immunohistochemistry was performed with the Vectra 3.0 Automated Quantitative Pathology Imaging System and the inFORM Cell Analysis software (Perkin Elmer, Waltham, MA, USA) including spectral separation of dyes, tissue, and cell segmentation, and automated cell counting of immune phenotypes.

Low-quality samples/images due to excessive IF background due to formalin fixation or loss of tissue integrity during the experimental procedures were discarded from the automated cell quantification analysis.

#### Vanderbilt validation data sets

Seventy melanoma and 8 lung cancer patient samples were procured based on the availability of tissue and were not collected according to a pre-specified power analysis (Additional file 3). Included in these, 42 melanoma samples and 7 lung cancer samples were baseline pre-anti-PD1 therapy. Remaining patients were treated with either anti-CTLA-4 alone or combinations of anti-PD-1 and anti-CTLA-4. Finally, 10 samples were obtained from progressing tumors in patients experiencing an initial response. Clinical characteristics and objective response data were obtained by retrospective review of the electronic medical record. Patients were classified in responders (complete response and partial response) and non-responders (progressive disease, mixed response, and stable disease) according to investigator assessed, RECIST defined responses. All patients provided informed written consent on IRB-approved protocols (Vanderbilt IRB # 030220 and 100178).

Total RNA quality was assessed using the 2200 TapeStation (Agilent). At least 20 ng of DNase-treated total RNA having at least 30% of the RNA fragments with a size >200 nt (DV200) was used to generate RNA Access libraries (Illumina) following the manufacturer's recommendations. Library quality was

assessed using the 2100 Bioanalyzer (Agilent), and libraries were quantitated using KAPA Library Quantification Kits (KAPA Biosystems). Pooled libraries were subjected to 75 bp paired-end sequencing according to the manufacturer's protocol (Illumina HiSeq3000). Bcl2fastq2 Conversion Software (Illumina) was used to generate de-multiplexed Fastq files.

For FOXP3, CD4, and CD8 IHC staining, slides were placed on a Leica Bond Max IHC stainer. All steps besides dehydration, clearing, and coverslipping were performed on the Bond Max. Heat-induced antigen retrieval was performed on the Bond Max using their Epitope Retrieval 2 solution for 20 min. Slides were incubated with anti-CD4 (PA0427, Leica, Buffalo Grove, IL), FOXP3 (14-4777-82, eBiosciences), or anti-CD8 (MS-457-R7, ThermoScientific, Kalamazoo, MI) for 1 h.

#### Analysis of IHC images with IHCCount

We considered 75 bright-field immunohistochemistry images from 33 melanoma patients and 16 images from 8 lung cancer patients (Vanderbilt cohorts). However, 3 melanoma patients had to be excluded from the analysis due to the low quality of the staining or poor tissue preservation. In total, we analyzed 72 images stained for CD4, CD8, and FoxP3 from 32 melanoma patients and 16 images stained for CD4 and CD8 from 8 lung cancer patients. To quantify both the number of total cells and tumor-infiltrating immune cells from the melanoma and lung cancer IHC images, we implemented a computational workflow, called IHCCount, using free open-source software tools. In this workflow different analytical tasks were performed, including image pre-processing, training of pixel classifiers, image segmentation, and analysis, together with cell counting and additional measurements of the tumor-covered area. The methodology of the analysis is described as follows.

To prepare the IHC images for further analysis, we used the script collection (bftools) from the consortium of Open Microscopy Environment (OME) [37]. First, the bright-field images were extracted as TIF files with the highest resolution from the image containers, available in Leica (SCN) format. Each of these high-resolution images (0.5  $\mu$ m/pixel,  $\times 20$  magnification) was then subdivided into equally sized, non-overlapping image tiles (2000  $\times$  2000 pixels) in order to limit the computational costs of the subsequent analytical tasks. The open-source software ilastik [38] and its "Pixel Classification" module were used to manually annotate objects of interest and generate classifiers that distinguish positively stained cells and nuclei from background and stromal tissue. For each sample, a set of 3 to 5 representative image tiles was randomly selected for training,

considering the diverse nature of the obtained images (caused, for instance, by the presence of artifacts, differences in illumination, and staining intensities). As a result, we obtained two classifiers, one to classify pixels belonging to positively stained cells and the other to classify pixels belonging to nuclei. In addition, both could classify background and stromal tissue. The classifiers were subsequently used in a batch process to obtain two sets of probability maps for each tile. Both sets were exported as multichannel TIF (32-bit float), where each channel represented the probabilities of one of the given classes (positively stained cells or nuclei, together with stromal tissue and background). Finally, we developed a Cellprofiler [39] pipeline (IHCCount.cppipe) that runs intensity-based operations to segment and identify positively stained cells, nuclei, and the area of total tissue using the previously generated probability maps together with the original image tiles as input files. The overall results for each image were obtained by summing up the results of the single image tiles.

All previously described steps of the analysis were implemented in a python script (runCP.py) and can be run from the command line. The pipeline, together with a description of the workflow, is publicly available at <https://github.com/mui-icbi/IHCCount>. IHCCount results for the Vanderbilt cohorts are reported in Additional file 3. Total cell densities per tumor sample to be used to scale quanTIseq cell fractions were estimated as the median number of nuclei per mm<sup>2</sup> across all images generated from that tumor.

IHCCount analysis of IHC images from CRC patients (Leiden cohort) was performed using the same approach adopted for the Vanderbilt cohorts.

### Benchmarking of deconvolution and marker-based methods

All methods were run in R using their original code or R package, except TIMER, which was run from the web interface (<https://cistrome.shinyapps.io/timer>). All methods were run with their default parameter settings. EPIC was run with the “BRef” signature on PBMC data and with the “Tref” signature on the tumor data. TIMER signatures for COAD, LUAD, and SKCM were used to analyze tumor data from CRC, lung, and melanoma patients, respectively; TIMER was not applied to PBMC data as the web interface only allows the analysis of tumor data. CIBERSORT estimates were aggregated across the major subtypes considered in the benchmarking (e.g., naïve and memory B cells were summed up to obtain total B cell estimates). For EPIC and xCell, T cell estimates were obtained by summing up CD4<sup>+</sup> and CD8<sup>+</sup> T cells. xCell “DC” scores were considered for dendritic cells, whereas the MCPcounter estimates from the “Monocytic lineage” were used to quantify monocytes.

### Computation of the deconvolution-based immunoscore and TB score from quanTIseq cell fractions

For the calculation of the deconvolution-derived immunoscore, we considered the fractions of CD8<sup>+</sup> T cells and CD3<sup>+</sup> T cells, where the latter was computed as the sum of CD8<sup>+</sup> T cell, CD4<sup>+</sup> T cell, and T<sub>reg</sub> cell fractions. CD3<sup>+</sup> and CD8<sup>+</sup> T cell fractions were dichotomized considering their median across all patients, computed separately for each cell type and cancer type, and used to identify two groups of patients: (1) “Lo-Lo” patients, with both CD3<sup>+</sup> and CD8<sup>+</sup> T cell fractions lower or equal to the median; (2) “Hi-Hi” patients, with both CD3<sup>+</sup> and CD8<sup>+</sup> T cell fractions higher than the median. The “Hi-Hi” and “Lo-Lo” classes for the T and B cell (TB score) were derived in an analogous manner, but considering the fractions of B cells and CD8<sup>+</sup> T cell estimated by quanTIseq.

### t-SNE plots

t-SNE plots of the TCGA solid cancers were generated with “Rtsne” R package. The t-SNE algorithm was run on the immune cell fractions estimated by quanTIseq, excluding the fraction of uncharacterized cells. We retrieved the annotation about microsatellite instability (MSI) from a recent paper [40], considering both the MSI categories of the TCGA consortium and the MSI/MSS classes predicted at a confidence level of 0.75. Unambiguous predictions were used to identify the MSI or MSS samples, whereas ambiguous predictions (MSI:1 and MSS:1), null predictions (MSI:0 and MSS:0), or unavailable samples were assigned to the “unknown” MSI state. Gene expression represented as *z* scores of log<sub>2</sub>(TPM+1). Before plotting, *z* scores higher than 3 (or lower than −3) were saturated to 3 (or −3).

### Statistical analysis

Correlation between numeric variables was assessed with Pearson’s correlation. The area under the receiver operating characteristic curve (AUROC) for multi-class classification was computed with the “multiclass.roc” function of the pROC R package. Constrained least squares regression was performed with the “lsei” function from the “limSolve” R package. The root-mean-squared error was computed as  $RMSE = \sqrt{\text{mean}((X_{\text{estimated}} - X_{\text{true}})^2)}$ . Statistically significant differences between two groups were tested with two-sided Wilcoxon’s test. For comparisons across multiple groups, Kruskal-Wallis test followed by two-sided Dunn’s pairwise post hoc was used. Normality of the data distribution was tested with Shapiro-Wilk test. Overall survival analyses were performed using the R package *survival* on TCGA survival data (“vital\_status”, “days\_to\_death”, and “days\_to\_last\_followup”). For each cancer type, patients were dichotomized in two groups according to the deconvolution-based immunoscore or TB score. The Kaplan-Meier estimator was used to generate survival



curves and logrank tests (corresponding to two sided  $z$  test) were applied.

## Results

### Development of quanTIseq deconvolution algorithm

We developed quanTIseq, a computational pipeline for the analysis of raw RNA-seq and tissue imaging data that quantifies the fractions and densities of ten different immune cell types relevant for cancer immunology (Fig. 1a). We first designed a novel signature matrix using RNA-seq data (Fig. 1b and Additional file 1). To this end, we collected a compendium of 51 publicly available RNA-seq data sets (Additional file 1) from ten different immune cell types: B cells, M1 and M2 macrophages, monocytes (Mono), neutrophils (Neu), natural killer (NK) cells, non-regulatory CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, T<sub>reg</sub> cells, and myeloid dendritic cells (DC). These data were integrated with additional large-scale data resources from immune and non-immune cells and used to select the signature genes with the highest specificity and discriminative power to construct the immune cell signature matrix (details in the “Methods” section).

We then developed a deconvolution algorithm to estimate the absolute proportions (i.e., cell fractions referred to the total cells in the sample under investigation) of ten different immune cell types from bulk RNA-seq data. quanTIseq performs deconvolution using constrained least squares regression [41] to force the cell fractions to be non-negative and their sum not to exceed 1. By allowing this sum to be lower than 1, quanTIseq estimates also the proportion of uncharacterized cells (referred to as “other” cells from here on), namely cells that are present in the cell mixture of interest but that are not represented in the signature matrix (e.g., cancer cells). After regression, quanTIseq normalizes the immune cell fractions by a scaling factor in order to correct for differences in total mRNA content per cell. The deconvolution of closely related cell types (e.g., T<sub>reg</sub> cells and non-regulatory CD4<sup>+</sup> T cells) is inherently hampered by the high correlation of their expression signatures (multicollinearity) and can result in the underestimation or “dropout” of low-abundance T<sub>reg</sub> cells [17]. As there is currently no consensus on whether regularization methods can overcome multicollinearity in regression-based deconvolution [42, 43], we adopted a heuristic strategy to specifically address the issue of T<sub>reg</sub> cell dropouts. Further details on quanTIseq algorithm are reported in the “Methods” section.

Deconvolution methods usually take as input a matrix summarizing the gene expression levels of the mixtures of interest [15] computed from raw expression data. These data can be profoundly different from the signature matrix used for deconvolution, both in terms of gene annotation and normalization of gene expression values. To avoid

issues arising from missing signature genes and different data-normalization procedures, quanTIseq implements a full pipeline for the analysis of raw RNA-seq data that builds the mixture matrix using the same approach employed for the signature matrix (described in the “Methods” section). The quanTIseq pipeline consists of three analytical steps, as depicted in Fig. 1c: (1) pre-processing of raw RNA-seq reads (single- or paired-ends) to remove adapter sequences, trim low-quality read ends, crop long reads to a maximum length, and remove short reads; (2) quantification of gene expression as transcripts per millions (TPM) [44]—which are suitable for expression deconvolution based on linear regression [45]—and raw counts; and (3) expression normalization, gene re-annotation, and deconvolution of cell fractions. A unique feature of quanTIseq is the possibility to perform *in silico* multiplexed immunoprofiling by complementing the deconvolution results with information from image analysis of IHC, IF, or H&E tissue slides. If total cell densities estimated from images are available, they are used by quanTIseq to scale the fractions of all the deconvoluted immune cell types to cell densities (step 3 in Fig. 1c).

quanTIseq was containerized using Docker (<https://www.docker.com>) and Singularity (<https://www.sylabs.io/singularity>) to simplify the installation and usage of all tools and dependencies, thereby standardizing data analysis and making it easily accessible by a broader audience. quanTIseq can be run on Mac OS X and Linux systems and is available at <http://icbi.at/quantiseq>.

### Validation of quanTIseq using simulated RNA-seq data and published data sets

To benchmark quanTIseq on well-defined cell mixtures, we simulated 1700 RNA-seq data sets of human breast tumors characterized by different immune infiltrate scenarios. The data were generated by mixing different proportions of RNA-seq reads from tumor and immune cells and by simulating different sequencing depths (details in the “Methods” section). In order to avoid the use of the same data set for the mixture and signature matrix in the benchmarking, we adopted a leave- $K$ -out cross-validation approach. Briefly, for each simulated mixture to be deconvoluted, a signature matrix was built excluding the  $K$  RNA-seq data sets included in the simulated mixture. quanTIseq obtained a high correlation between the true and the estimated fractions and accurately quantified tumor content, measured by the fraction of “other” cells (Additional file 2: Figure S1).

We then validated quanTIseq using experimental data from a previous study [46], in which peripheral blood mononuclear cell (PBMC) mixtures were subjected to both RNA-seq and flow cytometry. A high accuracy of the quanTIseq estimates was also observed for this data set (Fig. 1d and Additional file 2: Figure S2).

Additionally, we tested quanTIseq on two published microarray data sets used to validate previous deconvolution methods [17, 47]. Although quanTIseq is designed for RNA-seq data and might show lower accuracy on pre-computed expression data due to the lack of important signature genes and due to the different dynamic range of hybridization-based and RNA-seq technologies, it showed good deconvolution performance also on these data sets (Additional file 2: Figures S3 and S4).

We then applied quanTIseq to over 8000 TCGA samples across 19 solid malignancies. As no gold-standard measures were available for these samples, we considered previous estimates of lymphocytic infiltration [48] and tumor purity [24] available for a subset of the TCGA patients to further assess the validity of quanTIseq results. First, we compared the fraction of lymphocytes estimated by quanTIseq, computed by summing up the cell fractions of B cells, NK cells, CD4<sup>+</sup> and CD8<sup>+</sup> T cells, and T<sub>reg</sub> cells, with the “lymphocyte score”, a semi-quantitative measure of the number of tumor-infiltrating lymphocytes estimated previously from H&E-stained section slides of melanoma tumors ( $n = 468$ ) [48]. Although the two approaches were based on different features of the immune contexture, i.e., molecular vs. morphological, and sequencing data and images are usually generated from different tumor portions, their estimates showed a high agreement (Additional file 2: Figure S5a).

Second, we considered TCGA tumor purity values estimated in a previous work with a consensus approach integrating four computational methods based on RNA-seq, methylation, and mutational data [24]. We compared these purity values with the fraction of “other” cells inferred by quanTIseq for all cancer types for which both estimates were available for at least 100 patients. Although the fraction of “other” cells does not directly represent tumor purity as it can include different cell types (e.g., stromal cells), we reasoned that a large proportion of these cells are tumor cells and therefore a positive correlation between these two variables in solid tumors should be expected. Indeed, the fraction of “other” cells estimated by quanTIseq had a significant positive correlation with tumor purity in all cancer types, with a correlation ranging from 0.29 in glioblastoma (GBM) to 0.72 of skin cutaneous melanoma (SKCM) (Additional file 2: Figure S5b).

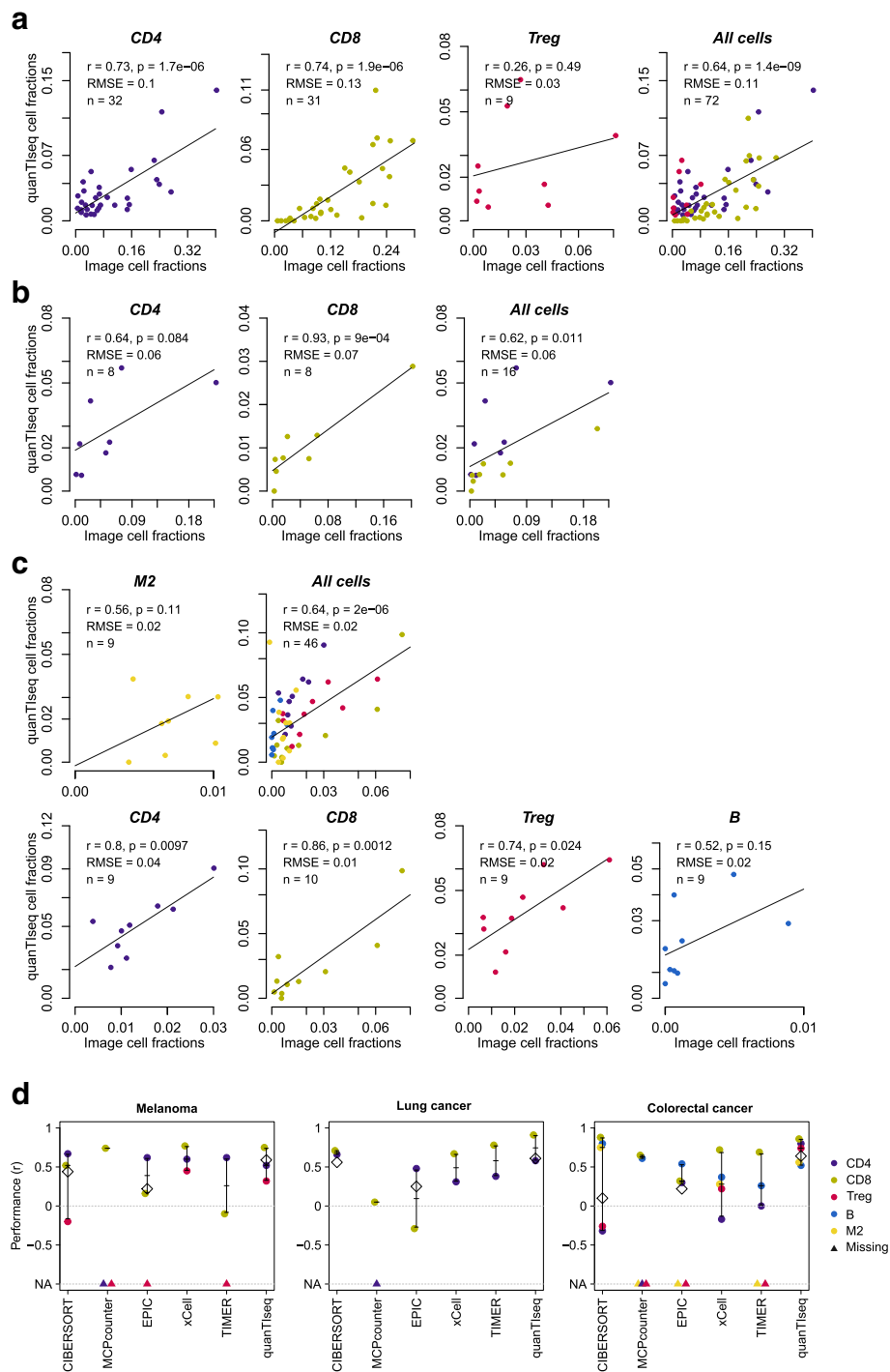
#### Validation of quanTIseq with flow cytometry immunoprofiling and IHC/IF data

As most of the validation data sets available in the literature are based on microarray data or consider a limited number of phenotypes, we generated RNA-seq and flow cytometry data from mixtures of circulating immune cells collected from nine healthy donors. The mixtures were generated by admixing low fractions of

polymorphonuclear (PMN) cells with PBMC extracted from the same donor samples (see the “Methods” section). Flow cytometry was used to quantify all the immune sub-populations considered by quanTIseq except macrophages, which are not present in blood. Comparison of quanTIseq estimates with the flow cytometry cell fractions showed a high correlation for all the single cell types (Fig. 1e and Additional file 2: Figure S6) and an overall correlation of 0.87. In particular, quanTIseq accurately quantified closely related cell types like non-regulatory CD4<sup>+</sup> T and T<sub>reg</sub> cells, as well as low-abundance dendritic cells (Additional file 2: Figure S6).

Finally, we validated quanTIseq using three independent cancer data sets (Additional file 2: Table S1). The first data set was generated from 70 tumor samples collected from melanoma patients. We carried out RNA-seq and, wherever possible, IHC staining for CD8<sup>+</sup>, CD4<sup>+</sup>, or FOXP3<sup>+</sup> cells from consecutive whole-tissue slides. To quantify specific immune cells from the scanned images, we developed an analysis pipeline (available at <https://github.com/mui-icbi/IHCCount>) to perform semi-automatic cell counting. The second data set was generated in an analogous manner using eight lung cancer samples and IHC images stained for CD8<sup>+</sup> and CD4<sup>+</sup> T cells. The third data set was generated from tumor samples of ten CRC patients. RNA-seq data, IF-stained slides for CD8<sup>+</sup> T cells and M2 macrophages (CD68<sup>+</sup>HLA-DR<sup>+</sup>CD163<sup>+</sup>), and IHC slides for CD4<sup>+</sup> T and T<sub>reg</sub> cells were generated and analyzed, wherever possible. Cell densities were then quantified with Perkin Elmer (<http://www.perkinelmer.com>) proprietary software for automated quantitative pathology (details in the “Methods” section). For all the three cancer cohorts, the cell fractions obtained with quanTIseq showed a good agreement with the IF/IHC-based estimates, computed both as cell fractions (i.e., ratio between positive cells and total nuclei) (Fig. 2a–c) and cell densities (positive cells per mm<sup>2</sup>) (Additional file 2: Figure S7). CD8<sup>+</sup> T cells were estimated robustly in all the three data sets ( $r = 0.74–0.86$ ,  $p \leq 0.0012$ ), whereas T<sub>reg</sub> cells, B cells, and M2 macrophages showed a lower agreement, with positive but non-significant correlations, likely due also to the small sample size and limited dynamic range of cell fractions. It is worth noting that these discrepancies might be also due to the different tumor portions used to generate images and RNA-seq data, as well as to the intrinsic limitation of using 1-to-3 cell markers for identifying distinct cell types from IHC/IF images.

We also used the IHC images from CRC patients’ samples to benchmark our IHCCount pipeline. We compared the cell fractions and densities obtained with IHCCount for CD4<sup>+</sup> T cells, T<sub>reg</sub> cells, and B cells with those obtained using Perkin Elmer (<http://www.perkinelmer.com>, details in the “Methods” section) proprietary software



**Fig. 2** (See legend on next page.)

(See figure on previous page.)

**Fig. 2** Validation of quanTIseq using tumor RNA-seq data and IF/IHC images. Comparison of quanTIseq cell fractions with those inferred for IF/IHC images from melanoma (**a**), lung cancer (**b**), and colorectal cancer (**c**) patients. Deconvolution performance was assessed with Pearson's correlation ( $r$ ) and root-mean-square error (RMSE) considering image cell fractions (ratio of positive cells to total nuclei) as ground truth. The line represents the linear fit. **d** Performance of quanTIseq and previous computational methods obtained on the three validation cohorts: melanoma, lung cancer, and colorectal cancer patients. Methods performance was quantified using Pearson's correlation ( $r$ ) considering image cell fractions as ground truth. Correlations for single cell types are displayed as dots, together with whiskers and horizontal bands representing median and 95% confidence intervals. Missing cell types are visualized as triangles at the bottom of the plots. Diamonds indicate the overall correlation obtained considering all cell types together; not shown for marker-based methods, which do not allow intra-sample comparison. B, B cells. CD4, total CD4<sup>+</sup> T cells (including also CD4<sup>+</sup> regulatory T cells); CD8, CD8<sup>+</sup> T cells; M2, M2 macrophages; T, Treg; regulatory T cells

for automated quantitative pathology—used here as gold standard for quanTIseq validation. The two approaches showed a high positive correlation both for cell fractions (Additional file 2: Figure S8a) and cell densities (Additional file 2: Figure S8b), although with a slight lower estimation of CD4<sup>+</sup> T and B cells for IHCCount.

Finally, we used the unique validation data set generated in this study also to compare quanTIseq performance with that of recent methods for the quantification of immune cells from expression data: CIBERSORT [17], MCPcounter [10], EPIC [19], xCell [11], and TIMER [18]; the latter was applied only to tumor data (details in the “Methods” section). Compared to deconvolution and marker-based methods, quanTIseq robustly obtained positive correlations across all cell types and data sets and scored amongst the top performers in all the assessments (Fig. 2d, Additional file 2: Figure S9 and Table S2). It is worth noting, however, that comparison of different deconvolution methods strongly depends on data type and pre-processing, on the number and type of immune cells considered (e.g., rare and similar cell types, considered by some methods but not by others, are more difficult to quantify), and on whether the estimates can be interpreted as cell fractions or not (see also a recent review [16]). Overall, our extensive benchmarking demonstrates the high accuracy and robustness of quanTIseq for quantification of immune cells from blood and tumor samples.

#### Activation of the CXCL9/CXCR3 axis is associated with immune infiltration in solid cancers

A comprehensive inventory of the molecular determinants that shape the tumor immune contexture has yet to be determined. In an attempt to identify promising candidates, we examined the association between the immune contexture and a set of features describing the genotypes of human cancers. For this purpose, we used quanTIseq to reconstruct the immune contexture of solid tumors from RNA-seq data of more than 8000 TCGA samples across 19 solid malignancies, and we assessed the correlation between absolute cell proportions and different genomic features: mutational load, neoantigen load, tumor heterogeneity, and fractions of mutations with clonal and subclonal origin. Surprisingly, there was either low or no correlation between these genetic correlates and the abundances of

tumor-infiltrating immune cells (Additional file 2: Figure S10). Moreover, the overall lymphocytic infiltration and the sum of all adaptive or innate immune cell fractions were only weakly associated with the mutational features in our pan-cancer and cancer-specific assessments.

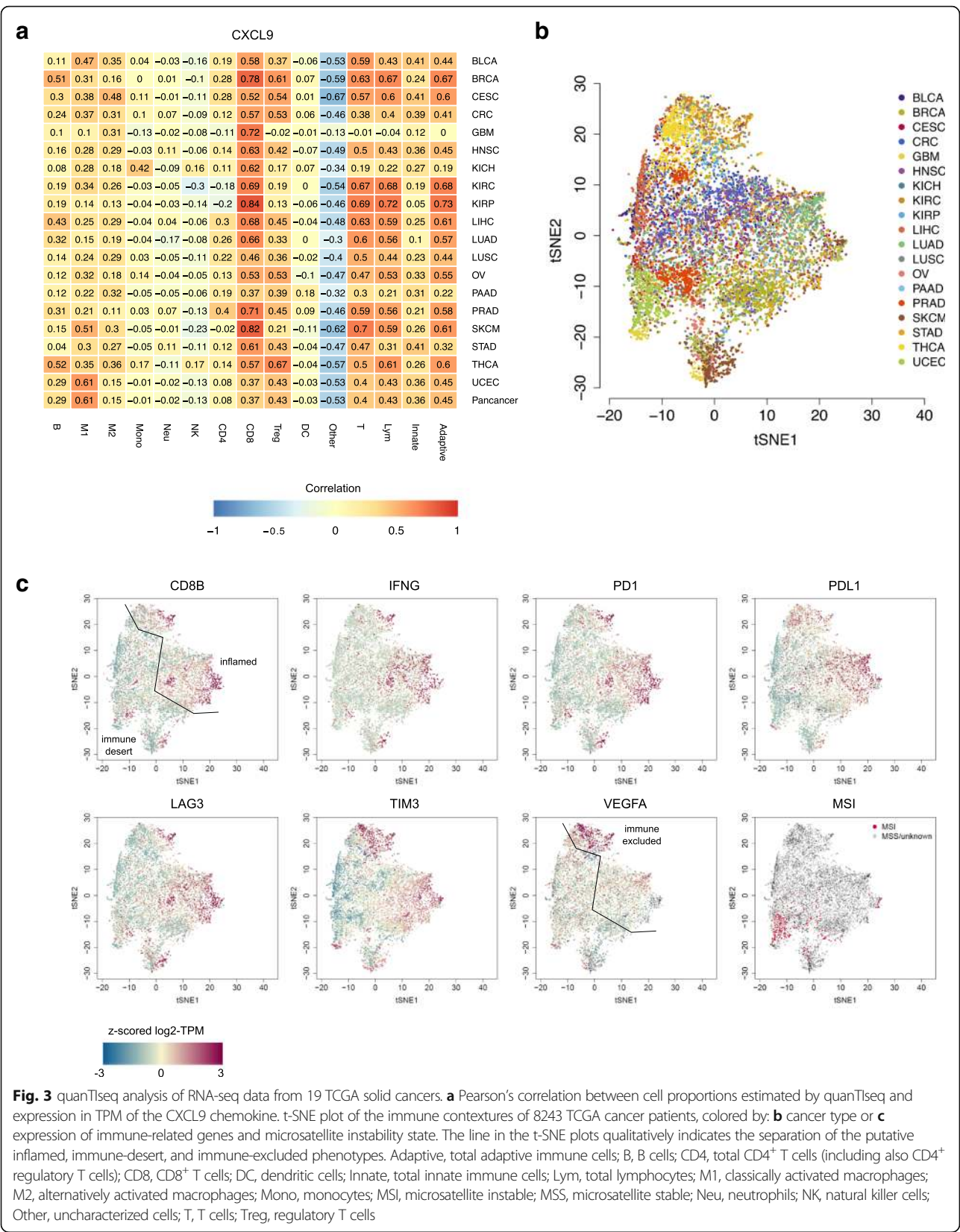
We have previously used biomolecular-network reconstruction to identify T cell homing factors associated with survival in CRC and pinpointed specific chemokines (CX3CL1, CXCL9, CXCL10) and adhesion molecules (ICAM1, VCAM1, MADCAM1) associated with high densities of intratumoral T cell subsets [49]. Therefore, we assessed the association between the expression of relevant chemokines, chemokine receptors, and adhesion molecules and the abundances of individual immune cell types (Additional file 2: Figure S11). We observed a high correlation between CD8<sup>+</sup> T cell fractions and the expression of CXCL9 chemokine (Fig. 3a) and chemokine receptor CXCR3 (Additional file 2: Figure S11b) and, for some cancer types, with CXCL10 expression (Additional file 2: Figure S11a). The CXCL9/CXCR3 axis regulates immune cell migration, differentiation, and activation and is therefore an important target for cancer therapy [50].

In summary, our results obtained using quanTIseq on bulk RNA-seq data from the TCGA suggests that the activation of the CXCR3/CXCL9 axis, rather than the genotype of the tumor, is associated with intratumoral cytotoxic T cells infiltration, and challenges the previous notion that the mutational burden is strongly associated with an increased infiltration of immune cells [51].

#### Pan-cancer analysis reveals highly heterogeneous immune contextures within and across solid cancers

We have previously shown that mutation and neoantigen profiles are highly heterogeneous on a sample by sample basis, being mostly characterized by passenger alterations that are rarely shared between patients [13]. However, despite this huge variability in their genotypes, tumors present common transcriptional signatures describing few molecular subtypes. For instance, analyses of a large number of samples identified four CRC subtypes with clear biological interpretability, called consensus molecular subtypes (CMS) [52]. Similarly, the immune profiles of human





cancers can be grouped into three major phenotypes, which are associated with response to PD1/PDL1 blockade: immune-inflamed, immune excluded, and immune desert [2]. Hence, we hypothesized that despite the genetic heterogeneity, human tumors converge to a limited number of immunological states quantified by the immune contextures. To test this hypothesis, we used dimensionality reduction based on the t-Distributed Stochastic Neighbor Embedding (t-SNE) [53] approach to visualize the 8243 immune contextures reconstructed by quanTIseq across 19 TCGA solid cancers (Fig. 3b and Additional file 2: Figure S12). Most of the cancer types did not create clearly distinct clusters, indicating highly heterogeneous immune contextures within and across cancers. Although some clusterization was visible for subsets of melanoma (SKCM), thyroid cancer (THCA), uterine cancer (UCEC), breast cancer (BRCA), and lung adenocarcinoma (LUAD) patients, a large overlap is seen for most of the cancer types. Visualization of gene expression (Fig. 3c) and immune cell fractions (Additional file 2: Figure S13) revealed two major clusters that might identify patients characterized by a high infiltration of cytotoxic CD8<sup>+</sup> T cells typical of the inflamed phenotype (right cluster in Fig. 3c with high CD8B expression), opposed to the immune-desert phenotype (left cluster in Fig. 3c with low CD8B expression) [2]. The inflamed phenotype was further associated with high expression of interferon gamma (IFNG), as well as with upregulation of immune checkpoints like PD1 and PDL1 and exhaustion markers like LAG3 and TIM3. Intriguingly, the plot also shows a cluster of patients characterized by high CD8B and VEGFA expression (top sub-cluster in Fig. 3c), which might correspond to an immune-excluded phenotype [2].

Based on the results of a recent clinical study [54], cancers with microsatellite instability (MSI) including CRC, uterine cancer, and ovarian cancer can be now treated with PD1 blockers. We therefore analyzed the immune contextures of MSI cancers from the TCGA cohorts (Fig. 3c). Similarly to the pan-cancer analyses, we found no distinct clusters also for this subgroup. Compared to their microsatellite stable (MSS) counterparts, MSI cancers were characterized by a significantly lower infiltration of M2 macrophages ( $p = 5.03 \cdot 10^{-8}$ ) and neutrophils ( $p = 1.28 \cdot 10^{-17}$ ) and by a significantly higher infiltration of M1 macrophages ( $p = 3.66 \cdot 10^{-3}$ ), NK cells ( $p = 5.76 \cdot 10^{-7}$ ), CD8<sup>+</sup> T cells ( $p = 1.75 \cdot 10^{-4}$ ), T<sub>reg</sub> cells ( $p = 1.34 \cdot 10^{-3}$ ), and dendritic cells ( $p = 3.67 \cdot 10^{-3}$ ).

In summary, we could show that, for human solid tumors, neither the classification according to the mutational load (MSI vs. MSS) nor the classification according to the anatomical site converges to a limited number of immunological states quantified by the immune contextures. However, it appears that some cancer subtypes exhibit similar immune contextures associated with specific genotypes as recently shown by us [13] and others [51].

### Deconvolution-based immune scores are associated with survival in solid cancers

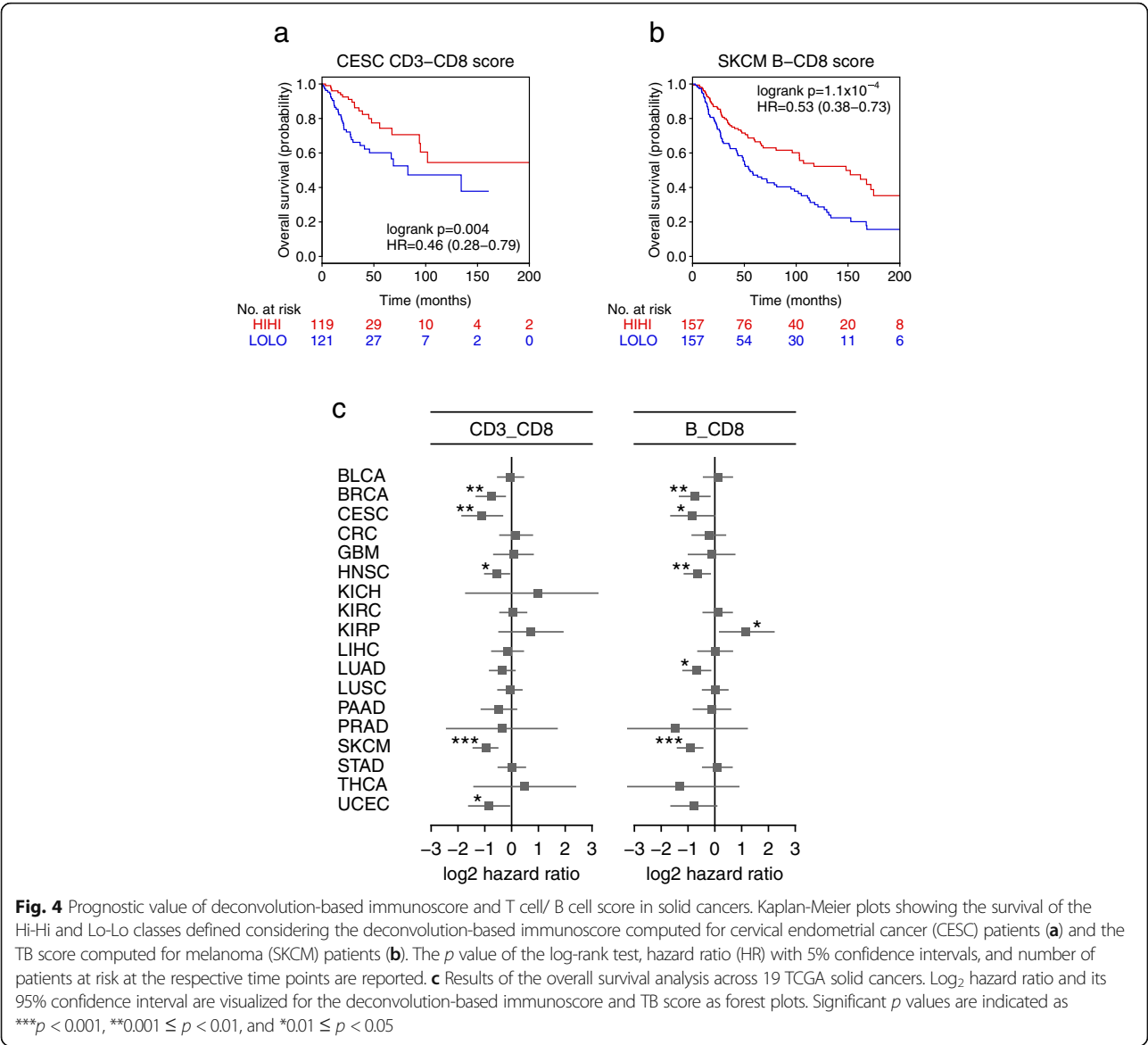
The immunoscore, a scoring system defined to quantify the immune infiltrates from tumor imaging data, has been demonstrated to be a prognostic marker superior to the TNM staging system in CRC [55]. The immunoscore is based on the enumeration of two lymphocyte populations (CD3<sup>+</sup> and CD8<sup>+</sup>) in the tumor core and invasive margin, and it can assume values from 0, when low densities of both cell types are found in both regions, to 4, when high densities are found in both regions. Recently, it was shown that the immunoscore and a newly proposed T and B cell score (TB score) were the strongest predictors of disease-free survival and overall survival in metastatic CRC [56].

We defined modified versions of the immunoscore and TB score based on the absolute fractions of the respective cell types deconvoluted by quanTIseq and tested their association with survival in solid cancers (see the “Methods” section). The results of the survival analysis using the computed TCGA cell fractions showed the prognostic value of the deconvolution-based immunoscore and TB cell score in five (BRCA, cervical squamous cell carcinoma [CESC], head and neck cancer [HNSC], SKCM, and UCEC) and six solid cancers (BRCA, CESC, HNSC, LUAD, and prostate adenocarcinoma [PRAD]), respectively (Fig. 4). The association was not significant for CRC as expected, due to the fact that spatial information of the immune cell distribution with respect to the tumor core and invasive margin could not be incorporated.

All quanTIseq results of the TCGA analysis have been deposited in The Cancer Immunome Atlas (<https://tcia.at>) [13] to make them available to the scientific community and facilitate the generation of testable hypotheses.

### Pharmacological modulation of the tumor immune contexture

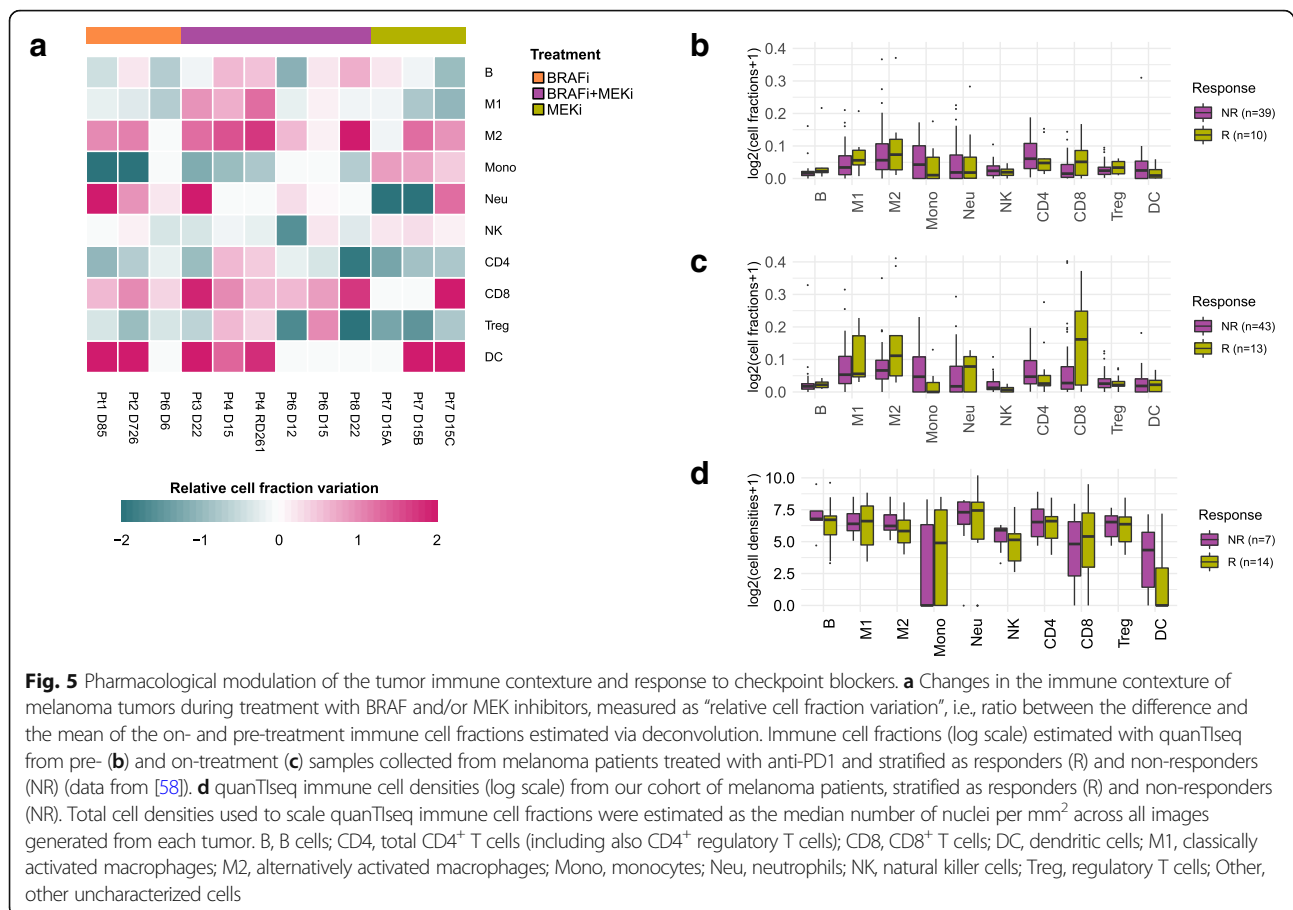
Beyond the extraction of prognostic markers, there is an urgent need to identify predictive markers for cancer immunotherapy with immune checkpoint blockers, as well as to determine the immunological effects of targeted agents [6]. We therefore used quanTIseq to investigate the pharmacological effects of targeted drugs on the immune contexture. We analyzed recently published RNA-seq data set from pre- and on-treatment tumor biopsies from seven melanoma patients treated with a BRAF inhibitors, MEK inhibitors, or a combination thereof [57]. quanTIseq deconvolution results showed large pharmacological remodeling of the immune contexture (Fig. 5a). Changes included a significant increase in dendritic cell fractions during treatment ( $p = 0.043$ ) and, to a lesser extent, an infiltration of CD8<sup>+</sup> T cells ( $p = 0.19$ ) and M2 macrophages ( $p = 0.07$ ). Thus, BRAF



and MEK inhibitors induce profound changes of the immune contexture. However, our analysis showed also patient-specific effects, further highlighting the need to develop immuno-oncology treatment strategies tailored to the individual patient.

Finally, in order to show the value of quanTIseq for informing cancer immunotherapy, we analyzed publicly available RNA-seq data from 51 pre- and 58 on-treatment samples collected from 65 melanoma patients treated with anti-PD1 [58]. quanTIseq analysis of pre- (Fig. 5b) and on-treatment samples (Fig. 5c) revealed higher B cell ( $p = 0.02$ ) and CD8<sup>+</sup> T cell ( $p = 0.03$ ) fractions, respectively, in responders compared to non-responders. Pre- and on-treatment samples from responder patients also showed higher M1 macrophage fractions, although the differences with non-responders were not statistically significant.

To further assess the predictive potential of quanTIseq, we considered 21 pre-treatment samples from our cohort of melanoma patients treated with anti-PD1 antibodies (nivolumab, pembrolizumab) and quantified the immune contexture using both bulk RNA-seq data and H&E-stained slides. We first carried out deconvolution using RNA-seq data and then scaled the fractions using total cell densities extracted from images to perform in silico multiplexed immunodetection. Total cell densities to be considered by quanTIseq can be computed from H&E-stained images (Fig. 1a). However, as H&E-stained images were not available for this cohort, we computed total cell densities as the median number of nuclei per mm<sup>2</sup> across all IHC images generated from a tumor. The cell densities estimated by quanTIseq showed a positive correlation with the image-derived densities



(Additional file 2: Figure S14). The deconvoluted cell densities of the ten immune cell types showed large heterogeneity across the patients and differences between responders and non-responders. For example, the densities of M1 macrophages as well as of CD4<sup>+</sup> and CD8<sup>+</sup> T cells were increased in responders compared to non-responders, although differences were not statistically significant ( $p > 0.09$ ), likely due to the limited number of samples (Fig. 5d). Further work with a larger number of samples is necessary to determine which immune cell type fractions or combined scores have predictive power for response to therapy with immune checkpoint blockers.

## Discussion

We developed quanTIseq, a computational pipeline for the analysis of raw RNA-seq and tissue imaging data that quantifies the absolute fractions and densities of ten different immune cell types relevant for cancer immunology. Unlike previous approaches, quanTIseq is specifically designed for RNA-seq data, which is the current reference technology for high-throughput quantification of gene expression [59]. To simplify data analysis and avoid inconsistencies between the mixture and the

signature matrix, we designed quanTIseq as a complete analytical pipeline that performs pre-processing of raw RNA-seq data, gene expression quantification and normalization, gene re-annotation, and estimation of cell fractions and densities. The results of our extensive validation using RNA-seq data from simulations, previous studies, blood cell mixtures, and three cancer patient cohorts demonstrate that quanTIseq can faithfully and quantitatively infer immune cell fractions from bulk RNA-seq data. Additionally, application of the method to publicly available data as well as data generated in this study revealed several important biological insights.

First, by analyzing more than 8000 TCGA samples, we showed that genomic features like mutational and neoantigen load, tumor heterogeneity, and proportion of clonal and subclonal mutations are only weakly associated with CD8<sup>+</sup> T cell fractions. In contrast, we found a stronger correlation between the activation of the CXCL9/CXCR3 axis and CD8<sup>+</sup> T cell infiltration in solid tumors, which would support the notion that CD8<sup>+</sup> T cells expressing the chemokine receptor CXCR3 can migrate into tumors following CXCL9 gradients [60]. This finding suggests that pharmacological modulation of the CXCL9/CXCR3 axis could be a therapeutic strategy to boost T cell recruitment,



thereby making also the immune-desert tumors [2] amenable to cancer immunotherapy. For instance, epigenetic reprogramming of genes expressing T helper ( $T_H$ )-1 chemokines like CXCL9 and CXCL11 might increase CD8<sup>+</sup> T cell infiltration into the tumor bed [60].

Second, our results indicate that the immune contexture is highly heterogeneous across and within solid cancers. This could partly explain the fact that the beneficial effects of cancer immunotherapy are observed only in a small fraction of patients. Furthermore, while the classification of common cancers into the three major immunophenotypes, namely immune inflamed, immune excluded, and immune desert, is conceptually appealing, it might not be sufficient to stratify the patients and thereby inform cancer immunotherapy. Our data suggest that the immune contexture and, hence, the immunophenotypes represent rather a continuous than a discrete variable, making it difficult to define cutoffs for precise stratification.

Third, the analysis with the deconvolution-based immunoscore and TB score supports the notion that combinations of different immunological features can have a stronger prognostic power than single markers. The lack of a significant prognostic value for some indications might be due to both, biological and technical reasons. For example, analyses of 10,000 samples showed remarkable degree of heterogeneity of the immune infiltrates across distinct organ-specific malignancies [51], suggesting that the cellular context is of utmost importance. Moreover, the high heterogeneity of the TCGA cohorts with respect to treatment and staging could be a possible confounding factor. Lastly, as we have previously shown that not only the density but also the spatial localization of the infiltrating immune cells plays a major role for the prognosis of tumor recurrence [3]. Enumeration of the immune cells in the core of the tumor and at the invasive margin markedly enhances the performance of the immunoscore. However, including this type of spatial information from the available TCGA images is challenging due to the limited performance of fully automated image analyses. Spatial lymphocytic patterns obtained using recent developments of deep learning tools [51, 61] might provide this missing information.

Fourth, quanTIseq analysis of the transcriptomes of patients treated with kinase inhibitors demonstrates profound pharmacological remodeling of the immune contexture. The immunological effects of conventional and targeted therapies came only recently into focus, fostering numerous clinical trials on combinatorial regimens of checkpoint blockers and targeted agents [62]. As bulk RNA-seq is now widely applied to profile fresh-frozen and archived tumor specimens, quanTIseq can be applied to effectively mine these data. Specifically, quanTIseq can be used to quantify the tumor immune contexture from large collections of formalin-fixed

paraffin-embedded (FFPE) samples in order to identify immunogenic effects of conventional and targeted drugs and hereby gain mechanistic rationale for the design of combination therapies.

Finally, our analysis of transcriptomics profiles from patients treated with anti-PD1 antibodies, although limited in sample size, shows the potential of quanTIseq for the extraction of immunological features that, alone or in combination, might predict the response to checkpoint blockade. Intriguingly, the higher infiltration of CD8<sup>+</sup> T cells in responder patients was not apparent from baseline samples but revealed itself shortly after the treatment start. This finding, also reported in a previous study on melanoma patients treated with CTLA4 and PD1 blockers [63], highlights the potential of early monitoring of the changes in the tumor immune contexture induced by checkpoint blockers. This could possibly reveal the mechanisms of resistance and enable identification of predictive markers for immunotherapy [64]. As more and more RNA-seq data sets from pre- and post-treatment samples of patients treated with checkpoint blockers will become available, we envision that quanTIseq will represent a useful resource to monitor the modulating effects of immunotherapy on the tumor immune contexture and extract candidate predictive markers.

We plan three lines of improvements of quanTIseq. First, as the transcriptomes of other non-malignant cell types from the tumor microenvironment will become available using bulk RNA-seq or single-cell RNA-seq, quanTIseq signature matrix can be extended to other cell types (e.g., cancer-associated fibroblasts) and optimized for specific cancer types. However, although immune cell phenotypes are known to depend on the specific tissue and disease context, to what extent expression signatures derived from the tumor microenvironment instead than from blood improve deconvolution performance remains to be clarified [19, 65, 66]. Second, spatial information on the localization of the infiltrating immune cells, i.e., localization in the center of the tumor and at the invasive margin, can be incorporated using annotation by a pathologist from images of H&E-stained slides. Finally, complementary information on the functional orientation of the infiltrating immune cells, including T cell anergy, exhaustion, or differentiation stage, can be derived from bulk RNA-seq data and included into the algorithm. However, since these functional states are not precisely defined in terms of unique expression signatures, a community-based consensus is required in order to include this type of information.

## Conclusions

In summary, we developed and thoroughly validated quanTIseq, a method for the quantification of the tumor

immune contexture using bulk RNA-seq data and histological images. Application of the tool to analyze thousands of samples from patients treated with conventional, targeted, or immunotherapeutic drugs revealed molecular and pharmacological modulators of the tumor immune contexture and immunological features underlying differential responses to immune checkpoint blockers. Hence, by analyzing carefully selected and well-annotated samples, our method holds promise to derive mechanistic rationale for the design of combination therapies and the development of predictive markers for immunotherapy. While quanTlseq represents an important contribution to the computational toolbox for dissecting tumor-immune cell interactions from RNA-seq data [15], we envision that it can be also applied to study autoimmune, inflammatory, and infectious diseases.

## Additional files

**Additional file 1:** quanTlseq signature matrix and used RNA-seq data from purified/enriched immune cells. RNA-seq data: cell type, study (full reference in the main text), Pubmed ID, Accession ID, Gene Expression Omnibus (GEO) ID of the sample, FASTQ ID, sequencing platform, read type, read length, link to the raw RNA-seq data. Signature matrix for the cell types: B cells (B.cells), classically activated macrophages (Macrophages.M1), alternatively activated macrophages (Macrophages.M2), monocytes (Monocytes), neutrophils (Neutrophils), natural killer cells (NK.cells), CD4+ T cells (T.cells.CD4), CD8+ T cells (T.cells.CD8), regulatory T cells (Tregs), and dendritic cells (Dendritic.cells). (XLSX 50 kb)

**Additional file 2: Figure S1.** In silico validation of quanTlseq. **Figure S2.** Validation of quanTlseq on PBMC RNA-seq data from [46]. **Figure S3.** Validation of quanTlseq on PBMC microarray data from. **Figure S4.** Validation of quanTlseq on PBMC microarray data from [47]. **Figure S5.** QuanTlseq analysis of RNA-seq data from TCGA tumors. **Figure S6.** Validation of quanTlseq on nine immune cell mixtures. **Figure S7.** Validation of quanTlseq in solid tumors using cell densities from IF/IHC images. **Figure S8.** Benchmarking of IHCCount on IHC images from CRC patients' samples. **Figure S9.** Performance of quanTlseq and previous deconvolution methods on PBMC data. **Figure S10.** Correlation between quanTlseq cell fractions and genetic variables. **Figure S11.** Correlation between quanTlseq cell fractions and expression of chemokines and adhesion molecules. **Figure S12.** t-SNE plots of 8243 TCGA samples colored according to cancer type. **Figure S13.** t-SNE plots of 8243 TCGA samples colored according to immune cell fractions. **Figure S14.** Validation of quanTlseq cell densities using IHC images. **Table S1.** Validation data considered in this study. **Table S2.** Performance of quanTlseq and previous deconvolution methods. **Table S3.** quanTlseq parameter settings. (PDF 25199 kb)

**Additional file 3:** Clinical and image data from the melanoma, lung cancer, and colorectal cancer cohorts. Clinical data: tumor identifier, immunotherapy, immune response (PD: Progressive Disease, MR: Marginal Response, SD: Stable Disease, CR: Complete Response, PR: Partial Response), sample type, and cancer type. Image analysis results obtained with IHCCount: cancer type, tumor identifier, marker gene, number of positively-stained cells, number of nuclei, tissue area in mm<sup>2</sup>, positive-cell densities (cells/mm<sup>2</sup>), total cell densities (cells/mm<sup>2</sup>), and positive cell fraction (positive/total). (XLSX 27 kb)

## Abbreviations

AUROC: Area under the receiver operating characteristic curve; BRCA: Breast invasive carcinoma; CCL: Cancer Cell Line Encyclopedia; CESC: Cervical squamous cell carcinoma; CGHub: Cancer genomics hub; CRC: Colorectal cancer; DC: Dendritic cells; GBM: Glioblastoma; GEO: Gene expression omnibus; GSEA: Gene set enrichment analysis; H&E: Hematoxylin and eosin; HNSC: Head and Neck squamous cell carcinoma; HR: Hazard ratio;

IF: Immunofluorescence; IHC: Immunohistochemistry; LUAD: Lung adenocarcinoma; M1: Classically activated macrophages; M2: Alternatively activated macrophages; Mono: Monocytes; MSI: Microsatellite instable; MSS: Microsatellite stable; Neu: Neutrophils; NK: Natural killer cells; NR: Non-responders; OME: Open microscopy environment; PBMC: Peripheral blood mononuclear cells; PMN: Polymorphonuclear cells; PRAD: Prostate adenocarcinoma; *r*: Pearson's correlation; R: Responders; RMSE: Root-mean-square error; RNA-seq: RNA sequencing; SKCM: Skin cutaneous melanoma; SRA: Sequence read archive; TB score: T and B cell score; TCGA: The Cancer Genome Atlas; T<sub>H</sub>: T helper cells; THCA: Thyroid cancer; TPM: Transcripts per millions; T<sub>reg</sub>: Regulatory T cells; UCEC: Uterine Corpus Endometrial Carcinoma

## Acknowledgements

We would like to thank Dr. Paul Hoertnagl (Innsbruck Blood Bank, Austria) for the collection of blood samples from healthy donors and Dr. Kristen L. Hoek (Vanderbilt Institute for Infection, Immunology and Inflammation, USA) for providing access to the flow cytometry data for algorithm validation. The results shown here are in part based upon data generated by TCGA Research Network (<https://cancergenome.nih.gov>).

## Funding

This work was supported by the European Union's Horizon 2020 research and innovation program under grant agreement No. 633592 (project "APERIM: Advanced Bioinformatics Tools for Personalised Cancer Immunotherapy"), by the Austrian Cancer Aid/Tyrol (project No. 17003, "quanTlseq: dissecting the immune contexture of human cancers"), by the Austrian Science Fund (FWF) (project No. T 974-B30 to FF, project I3291 to ZT, and project I3978 to ZT as member of the German Research Foundation (DFG) project TRR 241), by the Vienna Science and Technology Fund (Project LS16-025 to ZT), and by the European Research Council (advanced grant agreement No 786295 to ZT). The Leiden cohort work was supported by the Fight Colorectal Cancer-Michael's Mission-AACR Fellowship (2015) and Alpe d'HuZes/KWF Bas Mulder Award (UL2015-7664). The Vanderbilt cohort work was supported by the Vanderbilt-Incyte Research Alliance Program Grant (JMB, DBJ, YX), as well as R00CA181491 (JMB), K23CA204726 (DBJ), and the Breast Cancer Specialized Program of Research Excellence (SPOR) P50 CA098131.

## Availability of data and materials

The PBMC/PMN data set generated in this study is available from GEO (<https://www.ncbi.nlm.nih.gov/geo>) with accession GSE107572. The public data analyzed in this study are available on GEO with accessions GSE64655, GSE65133, GSE20300, GSE107572, and GSE91061. The data from colorectal cancer patients (Leiden cohort) are available from NdM on request. The data from melanoma and lung cancer patients (Vanderbilt cohorts) are available from JB on request. The datasets from the Leiden and Vanderbilt cohorts are part of larger studies and will be made available at later stage. A complete description of the data sets analyzed in this study is provided in Additional file 2: Table S1. All quanTlseq results from TCGA and from the patients treated with immune checkpoint blockers have been deposited in The Cancer Immunome Atlas (<https://tcia.at>) [13]. quanTlseq is available at <http://icbi.at/quantlseq>. IHCCount is available at <https://github.com/mui-icbi/IHCCount>.

## Authors' contributions

FF developed quanTlseq and performed deconvolution analyses and method benchmarking. FF, CP, and DR implemented quanTlseq software. FF and HH performed the statistical analyses. CM developed IHCCount. CM, DR, and ZL analyzed the Vanderbilt images. ZT, FF, AK, GL, WP, DW, and SS designed the validation experiment on blood mixtures. GL isolated the blood-derived immune cells and performed the flow cytometry experiment and data analysis. AK performed RNA extraction from the blood-derived immune cell mixtures. DJ contributed to the collection of Vanderbilt patient specimens and clinical annotation. YX and YW oversaw pre-processing and quality control of Vanderbilt RNA-seq data. MES oversaw histology techniques, analysis, and scoring techniques for the Vanderbilt cohorts. MVE and PEG performed IHC data pre-processing for the Vanderbilt cohorts. JB performed Vanderbilt data analysis and integration of techniques and performed and oversaw molecular assays and work involved in sample selection and isolation of nucleic acids. PC contributed to the project. MI, TPB, and NdM generated RNA-seq and multiplexed IF detection data for the Leiden cohort. FF and ZT wrote the manuscript. All authors read and approved the final manuscript.

### Ethics approval and consent to participate

**Leiden cohort.** This study was approved by the Medical Ethical Committee of the Leiden University Medical Center (protocol P15.282) and patients provided informed consent. Patient samples were anonymized and handled according to the medical ethical guidelines described in the Code of Conduct for Proper Secondary Use of Human Tissue of the Dutch Federation of Biomedical Scientific Societies. This research was conducted according to the recommendations outlined in the Helsinki declaration. **Vanderbilt cohorts.** All patients provided informed written consent. All human samples were obtained in accordance with the Declaration of Helsinki following protocols approved by Vanderbilt University Medical Center Institutional Review Board (Vanderbilt IRBs 030220 and 100178).

### Consent for publication

Not applicable.

### Competing interests

JB receives research support from Genentech/Roche, Bristol Myers Squibb, and Incyte Corporation; has received consulting/expert witness fees from Novartis; and is an inventor on provisional patents regarding immunotherapy targets and biomarkers in cancer. DJ is in the Advisory Board of Array, BMS, Genoptix, Incyte, Merck, and Novartis and has received research or travel support from BMS, Incyte, and Genentech. ZT is affiliated with ADIS and there are no competing interests related to this work. The remaining authors declare that they have no competing interests.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Biocenter, Division of Bioinformatics, Medical University of Innsbruck, Innrain 80, Innsbruck, Austria. <sup>2</sup>Division of Hygiene and Medical Microbiology, Medical University of Innsbruck, Innsbruck, Austria. <sup>3</sup>Department of Haematology and Oncology, Medical University of Innsbruck, Innsbruck, Austria. <sup>4</sup>Department of Pathology, Leiden University Medical Centre, Leiden, The Netherlands. <sup>5</sup>Vanderbilt University, Nashville, TN, USA. <sup>6</sup>Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. <sup>7</sup>Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, USA. <sup>8</sup>Department Pathology Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA. <sup>9</sup>Department of Medical Oncology and Internal Medicine VI, National Center for Tumor Diseases, University Hospital Heidelberg, Heidelberg, Germany. <sup>10</sup>Division of Translational Immunotherapy, German Cancer Research Center (DKFZ), Heidelberg, Germany. <sup>11</sup>Austrian Drug Screening Institute, Innrain 66A, Innsbruck, Austria.

Received: 26 October 2018 Accepted: 9 April 2019

Published online: 24 May 2019

### References

- Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer*. 2016;16:275–87.
- Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature*. 2017;541:321–30.
- Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*. 2006;313:1960–4.
- Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. *Nat Rev Clin Oncol*. 2017;14:717–34.
- Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer*. 2012;12:298–306.
- Galluzzi L, Buqué A, Kepp O, Zitvogel L, Kroemer G. Immunological effects of conventional chemotherapy and targeted anticancer agents. *Cancer Cell*. 2015;28:690–714.
- Tsujikawa T, Kumar S, Borkar RN, Azimi V, Thibault G, Chang YH, et al. Quantitative multiplex immunohistochemistry reveals myeloid-inflamed tumor-immune complexity associated with poor prognosis. *Cell Rep*. 2017;19:203–17.
- Petitprez F, Sun C-M, Lacroix L, Sautès-Fridman C, de Reyniès A, Fridman WH. Quantitative analyses of the tumor microenvironment composition and orientation in the era of precision medicine. *Front Oncol*. 2018;8:390.
- Finotello F, Eduati F. Multi-omics profiling of the tumor microenvironment: paving the way to precision Immuno-oncology. *Front Oncol*. 2018;8:430.
- Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016;17:218.
- Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. 2017;18:220.
- Angelova M, Charoentong P, Hackl H, Fischer ML, Snajder R, Krogsdam AM, et al. Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol*. 2015;16:64.
- Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18:248–62.
- Tappeiner E, Finotello F, Charoentong P, Mayer C, Rieder D, Trajanoski Z. TIminer: NGS data mining pipeline for cancer immunology and immunotherapy. *Bioinforma Oxf Engl*. 2017;33:3140–1.
- Hackl H, Charoentong P, Finotello F, Trajanoski Z. Computational genomics tools for dissecting tumour-immune cell interactions. *Nat Rev Genet*. 2016;17:441–58.
- Finotello F, Trajanoski Z. Quantifying tumor-infiltrating immune cells from transcriptomics data. *Cancer Immunol Immunother*. 2018;67:1031–40.
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12:453–7.
- Li B, Severson E, Pignon J-C, Zhao H, Li T, Novak J, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol*. 2016;17:174.
- Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *eLIFE*. 2017;6:e26476.
- Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;45:1113–20.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinforma Oxf Engl*. 2009;25:2078–9.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma Oxf Engl*. 2014;30:2114–20.
- Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34:525–7.
- Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. *Nat Commun*. 2015;6:8971.
- Birnbaum KD, Kussell E. Measuring cell identity in noisy biological systems. *Nucleic Acids Res*. 2011;39:9093–107.
- Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012;483:603–7.
- Allantaz F, Cheng DT, Bergauer T, Ravindran P, Rossier MF, Ebeling M, et al. Expression profiling of human immune cell subsets identifies miRNA-mRNA regulatory relationships correlated with cell type specific expression. *PLoS One*. 2012;7:e29979.
- Bonnal RJP, Ranzani V, Arrighi A, Curti S, Panzeri I, Gruarin P, et al. De novo transcriptome profiling of highly purified human lymphocytes primary cells. *Sci Data*. 2015;2:150051.
- Abbas AR, Wolslegel K, Seshasayee D, Modrusan Z, Clark HF. Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. *PLoS One*. 2009;4:e6098.
- Gong T, Hartmann N, Kohane IS, Brinkmann V, Staedtler F, Letzkus M, et al. Optimal deconvolution of transcriptional profiling data using quadratic programming with application to complex clinical blood samples. *PLoS One*. 2011;6:e27156.
- Liebner DA, Huang K, Parvin JD. MMAD: microarray microdissection with analysis of differences is a computational tool for deconvoluting cell type-specific contributions from tissue samples. *Bioinformatics*. 2014;30:682–9.

32. Marinov GK, Williams BA, McCue K, Schroth GP, Gertz J, Myers RM, et al. From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. *Genome Res.* 2014;24:496–510.
33. Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet.* 2013;29:569–74.
34. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinforma Oxf Engl.* 2007;23:1846–7.
35. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinforma Oxf Engl.* 2005;21:3439–40.
36. Ijsselssteijn ME, Brouwer TP, Abdulrahman Z, Reidy E, Ramalheiro A, Heeren AM, et al. Cancer immunophenotyping by seven-colour multispectral imaging without tyramide signal amplification. *J Pathol Clin Res.* 2019;5:3–11.
37. Linkert M, Rueden CT, Allan C, Burel J-M, Moore W, Patterson A, et al. Metadata matters: access to image data in the real world. *J Cell Biol.* 2010; 189:777–82.
38. Sommer C, Straehle C, Koethe U, Hamprecht FA. Ilastik: Interactive learning and segmentation toolkit. *Biomed Imaging Nano Macro 2011 IEEE Int Symp On. IEEE;* 2011. p. 230–233.
39. Kametsky L, Jones TR, Fraser A, Bray M-A, Logan DJ, Madden KL, et al. Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinforma Oxf Engl.* 2011;27:1179–80.
40. Cortes-Ciriano I, Lee S, Park W-Y, Kim T-M, Park PJ. A molecular portrait of microsatellite instability across multiple cancers. *Nat Commun.* 2017;8:15180.
41. Haskell KH, Hanson RJ. An algorithm for linear least squares problems with equality and nonnegativity constraints. *Math Program.* 1981;21:98–118.
42. Newman AM, Gentles AJ, Liu CL, Diehn M, Alizadeh AA. Data normalization considerations for digital tumor dissection. *Genome Biol.* 2017;18:128.
43. Li B, Liu JS, Liu XS. Revisit linear regression-based deconvolution methods for tumor gene expression data. *Genome Biol.* 2017;18:127.
44. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011;12:323.
45. Jin H, Wan Y-W, Liu Z. Comprehensive evaluation of RNA-seq quantification methods for linearity. *BMC Bioinformatics.* 2017;18:117.
46. Hoek KL, Samir P, Howard LM, Niu X, Prasad N, Galassie A, et al. A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. *PLoS One.* 2015;10:e0118528.
47. Shen-Orr SS, Tibshirani R, Khatri P, Bodian DL, Staedtler F, Perry NM, et al. Cell type-specific gene expression differences in complex tissues. *Nat Methods.* 2010;7:287–9.
48. Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. *Cell.* 2015;161:1681–96.
49. Mlecnik B, Tosolini M, Charoentong P, Kirilovsky A, Bindea G, Berger A, et al. Biomolecular network reconstruction identifies T-cell homing factors associated with survival in colorectal cancer. *Gastroenterology.* 2010;138:1429–40.
50. Tokunaga R, Zhang W, Naseem M, Puccini A, Berger MD, Soni S, et al. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - a target for novel cancer therapy. *Cancer Treat Rev.* 2018;63:40–7.
51. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang T-H, et al. The immune landscape of cancer. *Immunity.* 2018;48:812–30.e14.
52. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med.* 2015; 21:1350–6.
53. van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* 2008;9:2579–605.
54. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med.* 2015; 372:2509–20.
55. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, et al. Towards the introduction of the “Immunoscore” in the classification of malignant tumours. *J Pathol.* 2014;232:199–209.
56. Mlecnik B, Van den Eynde M, Bindea G, Church SE, Vasaturo A, Fredriksen T, et al. Comprehensive intrametastatic immune quantification and major impact of immunescore on survival. *J Natl Cancer Inst.* 2018;110:97–108.
57. Song C, Piva M, Sun L, Hong A, Moriceau G, Kong X, et al. Recurrent tumor cell-intrinsic and -extrinsic alterations during MAPK1-induced melanoma regression and early adaptation. *Cancer Discov.* 2017;7:1248–65.
58. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, et al. Tumor and microenvironment evolution during immunotherapy with Nivolumab. *Cell.* 2017;171:934–49 e16.
59. Finotello F, Di Camillo B. Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. *Brief Funct Genomics.* 2015; 14:130–42.
60. Nagarsheth N, Wicha MS, Zou W. Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. *Nat Rev Immunol.* 2017;17:559–72.
61. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep.* 2018;23:181–93 e7.
62. Schmidt C. The benefits of immunotherapy combinations. *Nature.* 2017;552: S67–9.
63. Chen P-L, Roh W, Reuben A, Cooper ZA, Spencer CN, Prieto PA, et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discov.* 2016;6:827–37.
64. Lesterhuis WJ, Bosco A, Millward MJ, Small M, Nowak AK, Lake RA. Dynamic versus static biomarkers in cancer immune checkpoint blockade: unravelling complexity. *Nat Rev Drug Discov.* 2017;16:264–72.
65. Schelker M, Feau S, Du J, Ranu N, Klipp E, MacBeath G, et al. Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nat Commun.* 2017;8:2032.
66. Sturm G, Finotello F, Petitprez F, Zhang JD, Baumbach J, Fridman WH, et al. Comprehensive evaluation of computational cell-type quantification methods for immuno-oncology. *bioRxiv.* 2019:463828. <https://doi.org/10.1101/463828>.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

